

Oklahoma State University

Economic Effects of Labor Unions

For the Fulfillment of the Honors College Degree in Economics

By

Henry Bibelheimer

Research Advisor: Bidisha Lahiri

Second Reader: Bill Mclean

Introduction:

I am going to be looking at how union participation affects state level outcomes. Labor unions might lower a state's GDP and raise poverty and unemployment, but inequality might go down because labor unions fight for the rights and benefits of employees on all levels. Hence my research will examine the data and discover the impact of labor unions on poverty, inequality, unemployment and GDP. I use various specifications to unearth the causation instead of just looking at the correlation observed in simple linear regressions.

Technique:

Simple regression:

The first calculation that I did is a simple regression that can be used to show correlation. This calculation is done with the equation below.

$$Y = \beta_0 + \beta_1 X$$

Y= Dependent variables (unemployment, poverty, income inequality, and GDP)

X= Labor union participation rate

Regression with Dummy Variables:

The next step in the process is to run a regression with dummy variables. In this regression we wanted each year and the state to have its own value. This is why I created a dummy for year and state, which allows for variation among the variables. We did this with a regression that looked similar to the equation below.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

X_2 = Dummy Variables (i.year, i.state)

When dummy variables are added to the regression it allows for a higher level of variation which shows a higher level of correlation than the simple regression.

Regression with interactive variables:

After doing a regression with dummy variables I added interaction to these variables. In this analysis I took each of the dummy variables and multiplied it to the union participation rate, which allows the effect of union participation to vary across states and years. The regression equation is the same as the last one, but this time interactive variables are inputted as well as dummy variables.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

X_2 = Dummy Variables

X_3 = Variables with interaction (i.year#c.unionp i.state_num#c.unionp)

Regression with lagged dependent variables:

First I had to create a lagging variable for each of the variables that we are studying. A lagging variable is a variable that has the value of the prior year. This was done with the command function below.

```
gen lag_poverty = poverty[_n-1]
```

To address any omitted variable issues lagging variables are introduced as explanatory variables. These lagged variables should embody the composite effect of any omitted variables. This calculation was done with the regression equation below.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

Y = Dependent variable

X_1 = Lagging dependent variable

X_2 = Labor union participation rate

X_3 = Lagging union participation rate

X_4 = Dummy variables

X_5 = Variables with interaction

Regression with First difference variables:

The final regression that I ran was a First Difference analysis. Before I could do this calculation I had to first create a variable that holds the value of the change in a variable in that year. This variable is created by subtracting the lagging value from the original value with the equation shown.

```
gen FD_poverty = (gdp - lag_poverty)
```

We then took this variable and added it to the regression equation below. The point of this calculation is to isolate the dependent variable, and show that the change in this variable is caused by the change in the independent variable.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Y = First difference dependent variable

X_1 = First difference labor union participation rate

X_2 = Dummy variables

X_3 = Interactive variables

Data:

We found data for labor union participation rate, unemployment, poverty, income inequality, and GDP. For every variable I found annual data for each state for the years 2000-2016 (except for income inequality because the GINI coefficient was not available for 2016 yet). Our unemployment data came from the Bureau of Labor Statistics online archive¹. I took advantage of the archive of data that the Census has for the percent of population that is below the poverty line². I used the Current Population Survey that covers information about Union Participation Percentages, which is calculated by the Bureau of Census for the Bureau of Labor Statistics³. I used information that Mark W. Frank compiled, using information from the IRS, for income inequality⁴. Our Gross Domestic Product information was pulled from the Bureau of Economic Analysis⁵. Here is some general information about the data that was collected and used for the regressions.

Variable	Obs	Mean	Std. Dev.	Min	max
Year	867	2008	4.9018	2000	2016
Unemployment	867	5.810035	1.9878	2.3	13.7
GDP (Millions)	867	279118.6	349782.4	17349	2602672
Poverty	867	12.7356	3.45083	4.5	25.8
Union Participation	867	11.3128	5.3809	1.6	26.1
Income Inequality	816	.59846	.036285	.52184	.711425

¹, data.bls.gov/pdq/SurveyOutputServlet

² www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html

³ data.bls.gov/pdq/SurveyOutputServlet

⁴ *Journal of Business Strategies*, vol. 31, no. 1, pages 241-263

⁵ www.bea.gov/regional/downloadzip.cfm

Results:

Effect of union Participation:

	Regression (i)	With state and year dummies (ii)	State and year dummies interacted with union participation (iii)	With lagged dependent variables (iv)	First differences (v)
Unemployment	0.0465194*** (0.000)	-0.0041362 (0.765)	-0.0761204 (0.496)	-0.073423 (0.525)	-0.0226877 (0.434)
GDP (Millions)	12739.15*** (0.000)	506.2754 (0.681)	-2334.966 (0.813)	-3443.24 (0.738)	2796.034 (0.245)
Poverty	-0.170933*** (0.000)	0.0225806 (0.344)	0.0538796 (0.290)	0.1394481 (0.469)	-0.0922696** (0.034)
Income Inequality	-0.000719*** (0.002)	-0.0011292*** (0.001)	-0.0021426 (-0.86)	-0.0019676 (0.434)	-0.00191*** (0.002)

P-values in parentheses, *** significant at 1% level, ** significant at 5% level

When the simple regression is run (column i), each of the four variables show a significant correlation on the 10%, 5%, and the 1% threshold. When analysis is done with state and year fixed effects (column ii) the data shows that a 1% increase in Union Participation rate is correlated with a 0.0011 decrease in the GINI coefficient of income inequality. With the First Difference regression (column v) we find that a 1% increase in the Union Participation rate causes a 0.0019 decrease in the GINI coefficient for income inequality. Both of these calculations are also significant on the 10%, 5%, and 1%. The relation does not emerge when the effect of union participation is allowed to vary with state and year (column iii) and when lagged dependent variables are used to capture omitted variables (column iv). Poverty has the only other statistically significant result, which was found when we ran our first difference regression. This calculation showed that a 1% increase in union participation causes a .0923 decrease in the poverty percentage, which is significant on the 10%, and 5%.

Conclusion:

I would like to go ahead and give you a little perspective on some of these numbers. A 1% increase in union participation causes a decrease of 0.0019 in the Gini coefficient (a standard

measurement for income inequality). For comparison the average GINI coefficient in 2015 was 0.609146, which means that a 1% increase in union participation will cause about a 0.312% decrease in income inequality.

Similarly a 1% increase in union participation causes a .0923 increase in this poverty percentage. The average percentage of a state that was living in poverty in 2016 was 12.5%. This means that a 1% increase in labor union participation will cause a 0.74% decrease in the percent of our population living in poverty.

Overall it is not surprising that all of the variables are correlated with the simple regression, because the variables that I am testing are all economic variables that fluctuate with the economy. It is surprising that Income Inequality and poverty are the only dependent variable with which Union Participation percentages have a causal relationship. This is because there are a lot of economic arguments that say Labor Unions have a negative relationship with GDP, and a positive relationship with unemployment and poverty. My findings show that an increase in Labor Union Participation will cause a decrease in the Gini coefficient, and a decrease in the percent of the population suffering from Poverty. Further research is necessary because these calculations only include the last 17 years of information.

Bibliography

“Income Inequality.” *Inequality.org*, inequality.org/facts/income-inequality/.

“Local Area Unemployment Statistics.” *U.S. Bureau of Labor Statistics*, U.S. Bureau of Labor Statistics, data.bls.gov/pdq/SurveyOutputServlet.

“Union Affiliation Data from the Current Population Survey.” *U.S. Bureau of Labor Statistics*, U.S. Bureau of Labor Statistics, data.bls.gov/pdq/SurveyOutputServlet.

US Census Bureau. “Historical Poverty Tables: People and Families - 1959 to 2016.” *U.S. Trade with Haiti*, www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html.

US Department of Commerce, et al. “Annual Gross Domestic Product (GDP) by State.” *U.S. Bureau of Economic Analysis (BEA)*, www.bea.gov/regional/downloadzip.cfm.

Frank, Mark. W. 2014 "A New State-Level Panel of Annual Inequality Measures over the Period 1916 - 2005" *Journal of Business Strategies*, vol. 31, no. 1, pages 241-263